

24 JUN 1983

A Discussion Paper
for a
New Approach to a Performance Appraisal Program

Purpose: To address the apparent failure of the recently revised PAR to work effectively, and to test whether a "new" system will gain the support and confidence of the workforce.

Discussion:

1. Perhaps the most persistent and intractable problem inherent to any performance appraisal exercise is its reliability as a measure of how well those evaluated, in fact, met the requirements of the job. It long has been recognized that performance appraisal is, for all practical purposes, a process which involves "rendering a judgment," where the criteria used are known largely by the individual(s) responsible. Although objectivity always is emphasized, few performance appraisals can be made without an element of subjectivity. This is where the rub comes.

2. Few people really accept the validity of performance appraisal because of the conviction that subjective judgments are filled with personal bias; i.e., each person has his or her own private frame of reference when evaluating others. In some instances, of course, the individual may not be consciously aware of the primary factors influencing the evaluation made. Many psychological variables may attend a given "rater-ratee" relationship but suffice it to say that the well-known axiom, "one can only hope to recognize and control bias not eliminate it," pertains. This is one of the chief reasons that seniority and tenure continue to be viewed by large numbers of workers as a fairer means of making personnel decisions dealing with matters of pay and retention.

3. The necessity of some form of performance appraisal is rather obvious, and there is little question as to its growing importance to employees and to personnel management generally. On the other hand, unless employee confidence supports the PAR program it can cause as many problems as it hopes to solve. Supervisory officials too must be satisfied that the performance appraisal process serves their needs and is not overly burdensome.

4. One thing is certain, experience has shown that performance appraisal does not work very well as a multi-purpose program. No performance appraisal system yet devised has succeeded simultaneously in (1) assessing the employee's worth to an organization; (2) guiding, developing, and highlighting the employee's career opportunities; and (3) motivating the employee to a sustained and stronger work performance. Although all are bona fide objectives, they are not likely to be achieved by a single performance appraisal design or system.

5. The design potential of the PAR form itself, of course, is almost limitless. The evidence is clear and persuasive that forms, per se, have little influence on whether an appraisal system works, but the appraisal process does have a significant impact on results. Consequently, the performance appraisal process should be developed and tested with care.

6. The Agency's PAR is a vital part of its personnel evaluation system. Many employees, maybe most, believe it is the only part. They consider the PAR to be the paramount factor influencing their destinies in the Agency. Employee apprehension increases measurably when Agency management takes action, for example, to try to lower the average PAR rating level over a concern for inflated rating levels. It is interesting to note that the PAR survey of 1980-81 showed that nearly 70 percent of the respondents are of the opinion that supervisors give subordinates higher ratings than they deserve, while a similar number believe the PAR accurately reflects the level of their own performance. In other words, the employee says, "Others get inflated ratings, I don't!" This is the second rub; employees support the downward adjustment of others performance appraisal level, but not their own. While employees will never complain about getting an inflated evaluation themselves, they will lose confidence in a PAR program they perceive as benefiting those receiving such inflated ratings. As long as promotion rates (and other positive consequences of high PAR ratings) are perceived to be at an acceptable level, i.e., when the "bennies" appear to be shared equitably, employees express satisfaction with the system. The converse is true no matter how accurate or well-designed the PAR happens to be.

7. Accepting the probability that no single performance appraisal design will prove to be totally satisfactory as a management tool, there are possibilities for improving the Agency's present system. For one thing the utility of individual task (key element) ratings is questionable. These ratings have no specific regulatory significance as they do in the competitive civil service, and do not serve as the sole basis for determining the PAR's overall rating level.

8. In addition, these individual task ratings also require the development of performance standards for each key element cited. This is viewed by supervisors as a grueling and difficult process at best, particularly if one is intent on identifying standards that satisfy the interests of both management and the employee. The concept and use of performance standards has posed continuing problems for Agency supervisors, and they would welcome any reduction in the need for them. I suggest that as an alternative we concentrate on the development of a single performance standard representing the job in its entirety. The development of a single standard for the job not only would be a less onerous task but likely would be more meaningful to supervisors and career panels, than the present system. Thus, a part of the "reform" of the present PAR program would entail modification of the present version of the Advance Work Plan (AWP).

9. Assuming that a revised PAR calling for only a single overall rating level would be acceptable, the next question leads to determining how many rating levels should be established from which this single rating would be drawn. There is considerable disagreement among investigators as to the answer to this question. The range of levels most often cited varies from three to seven. There are advocates supporting any number of levels within that range. No studies, to my knowledge, have demonstrated or given convincing evidence that one number of levels is clearly superior to another. In effect, the choice borders on being arbitrary. Certainly the arguments offered are not of equal strength, but the number of levels chosen will relate more to the predilections or concerns of those making the decision than to any scientific finding.

10. Attached are examples of different numbers of rating levels within the range discussed. Some of the pros and cons for each are affixed. It is important in reviewing these examples that several points be kept in mind as follows:

a. These are unfinished products, i.e., the wording in each of the performance categories is not necessarily in final form (applies both to the "performance descriptor" and its description). Word selection and usage in any evaluation measure are difficult, of course, since word choices are seldom non-controversial. I have listed various combinations of "performance descriptors" to demonstrate the subtle distinctions always to be found when word choices are made in communicating concepts. These various terms illustrate the differences in flavor, nuance, or perspective that results.

b. The term "performance descriptor" replaces the term rating level because the latter tends to direct evaluators toward the notion that the one being rated is to be placed on a rung (frequently predetermined) of a verbally structured ladder. The use of numbers also tends to foster a ladder effect. Whereas, the raters' attention always should focus first on describing the subordinate's work performance, as it is perceived by the rater, in relation to the performance standard for the job. Then and only then should the rater seek the performance descriptor and its "standardized description" which most nearly characterizes the subordinate's work performance.

c. The column which reflects the expected percentage of the employee population to which each performance category will apply is not based on any statistical concept or mathematical formula. It is an estimate arrived at somewhat intuitively (based on my perceptions of the Agency workforce) and is not cast in concrete. The figure represents in my judgment, the appropriate percentage of persons likely to be evaluated at the level indicated provided (1) that well-developed performance standards are used as a reference and (2) that supervisors are responsible in applying them. In effect, they represent "ball-park" figures that would be realized if the PAR program is ideally administered.

11. Another aspect of performance appraisal which some research studies suggest impacts favorably on performance improvement deals with the input employees can provide directly. It is a widely held belief that employees are likely to become more accepting and confident of the worth of performance appraisal as a valid tool for management if they feel personally involved in the process. Research studies tell us that subordinates who are asked to contribute to the appraisal process itself, by compiling information and providing it to supervisors before PARs are prepared, acquire a sense of participation and partial ownership of the program. This in turn effects performance improvement. Researchers conclude that this procedural strategy is successful because it offers... "concrete proof that the organization desires the subordinate's participation."¹

12. The only subordinate participation formally encouraged by the Agency is in the development of the key job elements and performance standards recorded in the AWP. Unfortunately, our survey results indicate that this is not a common practice. Except for a brief effort several years ago by the then ORPA, the Agency has never pursued nor formally tested a PAR program providing for direct subordinate input into the actual performance appraisal report itself. (The ORPA effort was terminated but unfortunately there is no documentation of the results; an assumption would have to be made that it was not successful.) The subordinate's role in the Agency has been essentially passive; i.e., the individual is simply informed after the fact of the supervisor's rating and comments. Some supervisors offer a draft copy of the PAR to the employee to read before it is made final, but I would argue that this is more apt to be a gesture to preclude a possible unpleasant confrontation than a sincere elicitation for meaningful input.

13. If we accept the premises that employee involvement is important and that employees generally know what their job demands of them and what it takes to perform well, then it follows that more employee participation, both in the development of the AWP, and in the preparation of the PAR itself, should have a positive effect on the whole process. Of great importance, of course, is the development of performance standards which, after all, represent the basic foundation of relative objectivity upon which successful performance appraisal rests. The exact nature of employee involvement in the total process would need to be worked out but their input into the evaluation process offers another potentially promising means of improving the Agency's current performance appraisal system.

¹Susan M. Resnick and Allan M. Mohrman, "The Design of Performance Appraisal Systems: Some Implications from Research Findings" Center for Effective Organizations Graduate School of Business Administration, University of Southern California.

14. In summary, to revitalize the effectiveness of the Agency's current performance appraisal program, I recommend the following:

- a. elimination of specific ratings for individual key elements;
- b. provision of employee input to the performance appraisal report before a finished copy of the PAR is prepared, with the exact nature of that input yet to be determined;
- c. adoption of a new rating system using overall performance descriptors in lieu of the current numbers.

Thurs May 28 1983

ROUTING AND RECORD SHEET

SUBJECT: (Optional) A Discussion Paper for a New Approach to a Performance Appraisal Program

FROM: <input type="text"/>	EXTENSION <input type="text"/>	NO. <input type="text"/>
DD/PA&E/OP 1006 Ames	<input type="text"/>	DATE 28 June 1983

STAT
STAT

TO: (Officer designation, room number, and building)	DATE		OFFICER'S INITIALS	COMMENTS (Number each comment to show from whom to whom. Draw a line across column after each comment.)
	RECEIVED	FORWARDED		


1. EA/D/OP 5E58 Hqs.	<i>JUN 1983</i>		<i>PLC</i>	Attached is a discussion paper prepared by <input type="text"/> concerning the performance appraisal system and a possible alternative approach. We'd like to discuss it with you at your convenience.
2. DD/OP <i>28 JUN 1983</i>			<i>12</i>	
3. D/OP				
4.				

STAT

ILLEGIB
STAT

5.				 <i>New Proposal</i> <i>Please see para 8.</i>
6.				
7.				
8.				
9.				

ILLEGIB

10.				 <i>RS - sky mty</i>
11.				
12.				
13.				
14.				
15.				

THREE PERFORMANCE CATEGORIES

Pro -

- minimizes the discrimination in performance judgments required of the supervisor.
- is the most basic method for certifying an employee's status for record purposes.
- recognizes that the large proportion of employees are in a performance range where no special personnel action is necessarily required or warranted.

Con -

- places the bulk of the workforce in a category too broad to be very useful.
- places considerable pressure on supervisors to inflate evaluations.
- could demoralize many employees; the category descriptor "satisfactory" tends to be used perjoratively in the Agency.

(Three Performance Categories)

Performance Descriptor	Expected % of Population	Description
Unsatisfactory	.5-2%	Performance does not meet the established work standard for the position and specifically demonstrates the individual's failure to meet one or more important job requirements.
*Satisfactory	80-85%	Performance meets the established work standard for the position. Work behavior may range from barely satisfactory to almost exceptional. Although the range of performance may vary considerably among the total population, the large majority would be expected to display a good knowledge of the job; would be dependable and resourceful in producing a quality product or service; and have good work relationships.
Exceptional	15-20%	Performance <u>far</u> exceeds the requirements of the <u>job</u> and may approximate a maximum possible level of achievement. The individual's impact on the job and the quantity and quality of his or her work product or service would be distinctly superior in comparison to others, and warrant special recognition.

*Performance (Work) Standard

FIVE PERFORMANCE CATEGORIES

Pro -

- allows one more level (in each direction) to better discriminate between the two extremes of work behavior.
- provides a middle level where the majority of employees are assumed to be best represented.
- represents the more popular model because the number of levels is believed to be neither too few nor too many.

Con -

- has a middle level which most employees currently view as an anathema.
- viewed by some as not permitting a sufficient spread to properly categorize performance variations.
- it's use in the past might predispose evaluators to follow "old habit patterns" and make it difficult for them to not inflate their ratings.

(Five Performance Categories)

Performance Descriptor	Expected % of Agency Population	Description
Unacceptable	.5-2%	Performance fails to meet minimal work requirements. Work behavior may be characterized by inadequate knowledge; an unsatisfactory display of the skills and abilities essential to the job; poor work relationships.
Adequate	10-20%	Performance is essentially acceptable having no crucial short comings. Work behavior may be characterized by periodic remissness in attending to duties and responsibilities; occasional variability in the quality or quantity of work output; some problems in working relationships.
*Competent	40-60%	Performance satisfies all job requirements and fully meets management's expectations. Work behavior ordinarily characterized by a quality and level of productivity that warrants no criticism; harmonious working relationships.
Noteworthy	15-25%	Performance clearly exceeds job requirements and is of a higher calibre than normally anticipated. Work behavior may be characterized by exceptional achievements in both the quantity and quality of work output; innovative work accomplishments.
Superlative	5-10%	Performance is conspicuously deserving of special recognition. Work behavior is characterized by highly impressive achievements normally accomplished by few persons in the performance of their duties.

*Performance (Work) Standard

SIX PERFORMANCE CATEGORIES

Pro -

- eliminates the middle category which most employees have an aversion to.
- places the bulk of the population in the upper (most attractive) half of performance categories.
- geared more to "positive" evaluations which evaluators are prone to make.

Con -

- stretches the credibility of the scale when only one performance standard is used.
- compresses the number of top level evaluations.
- makes efforts to achieve substantial increases in work quantity or quality more difficult. Tends to accept the notion that the large majority of employees are doing "more than is expected."

(Six Performance Categories)

Performance Descriptor	Expected % of Population	Description
Unacceptable	.5-2%	Performance at this level does not meet minimal requirements of the position. Work behavior may be characterized by inadequate job knowledge; failure to complete assigned work; excessive and serious errors in work produced; failure to get along with associates when this is essential to work accomplishments.
Passable	3-5%	Performance is at or slightly above the minimal requirements of the position. Work behavior may be characterized by inordinate slowness in getting assigned work accomplished; slipshod work products; mediocre work relationships.
Acceptable	15-25%	Performance approaches the level expected and generally meets the demands of the job.
*Commendable	40-60%	Performance exceeds, by a notch or two, the acceptable level. Work behavior may be characterized by a highly competent display of skills or abilities in producing work products; tasks or assignments are handled or accomplished in a manner which more than satisfies job requirements.
Praiseworthy	15-25%	Performance clearly exceeds the commendable level. Work behavior may be characterized by extraordinary ability to solve work problems or improve the units (section, branch, division, etc) operations and/or productivity.
Distinguished	3-5%	Performance <u>far</u> exceeds work requirements. Work behavior may be characterized by outstanding achievement and effectiveness seldom seen in any organization.
*Performance (Work) Standard		

SEVEN PERFORMANCE CATEGORIES

Pro -

- retains broad range of performance categories.
- offers a slightly "positively skewed" distribution which should appear more attractive to employees.
- represents an upgrading of the current PAR.

Con -

- too similar to current PAR. Most likely could not be introduced "uncontaminated."
- too many levels to use successfully with one performance standard.
- suffers from the use of a "center" category which many employees view as uncomplimentary.

(Seven Performance Categories)

Performance Descriptor	Expected % of Agency Population	Description
Unsatisfactory	.5-2%	Performance does not meet the established work standard and specifically demonstrates the individual's failure to meet job requirements. Work behavior may be characterized by failure to complete work; inability to avoid serious work errors; significant interpersonal conflicts on the job.
Marginal	2-4%	Performance reflects a serious work related problem. Work behavior characterized by marked deficiencies in providing work products or services.
Acceptable	10-20%	Performance is not faulty in any major respect. Work behavior may be characterized by an occasional drop in the quality or quantity of work produced; minor shortcomings may be observed.
*Satisfactory	30-50%	Performance fully meets all requirements and expectations of the job. Work behavior may be characterized by consistently high productivity, effective interpersonal relations; a generally exemplary performance.
Good	15-20%	Performance is more than satisfactory and better than expected. Work behavior may be characterized by the achievement of more than ordinary goals; shows obvious diligence in meeting assigned responsibilities and a willingness to take a "little extra time or effort" to assure the job is done "right."
Excellent	10-15%	Performance reveals an extraordinary level of achievement of work objectives. Work behavior is characterized by a rather dramatic display of ability and accomplishments readily noted by others.
Outstanding	3-5%	Performance is highly exceptional and not often observed. Work behavior is characterized by unusual proficiency and adroitness; a masterful job, well-done.

*Performance (Work) Standard